



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/804,516	03/19/2004	Arun Kwangil Iyengar	YOR920040025US1	7509
7590 Ryan, Mason & Lewis, LLP 90 Forest Avenue Locust Valley, NY 11560				
09/16/2009				
EXAMINER				
PHUNG, LUAT				
ART UNIT		PAPER NUMBER		
2416				
MAIL DATE		DELIVERY MODE		
09/16/2009		PAPER		

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary

Application No.

10/804,516

Applicant(s)

IYENGAR ET AL.

Examiner

LUAT PHUNG

Art Unit

2416

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 03 April 2009.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-25 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-25 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. _____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☐ Information Disclosure Statement(s) (PTO/SF/ICE)
Paper No(s)/Mail Date _____
- 4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date _____
- 5) ☐ Notice of Informal Patent Application
- 6) ☐ Other: _____

Response to Amendment

1. Applicant's arguments, see pages 1-4, filed April 3, 2009 with respect to the rejection(s) of claim(s) 1-25 under 103(a) have been fully considered and are persuasive. Therefore, the rejection has been withdrawn. However, upon further consideration, a new ground(s) of rejection is made in view of additional prior art references Mangipudi, et al. and Chen, et al.

Claim Rejections - 35 USC § 101

7. 35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

2. Claims 1-13 and 18-25 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter. Claims 1-13 and 18-25 fail to recite any structural tie to any class of invention and therefore do not satisfy the threshold tie to be eligible for patent protection under 35 U.S.C. 101. While the claim(s) recites a series of steps or acts to be performed, a statutory "process" under 35 U.S.C. 101 must (1) be tied to another statutory category (such as a particular apparatus), or (2) transform underlying subject matter (such as an article or material) to a different state or thing. The instant claims neither transform underlying subject matter nor positively tie to another statutory category that accomplishes the claimed method steps and, therefore, do not qualify as a statutory process. In particular, the methods include the steps which appear purely directed to mental steps or mathematical manipulations of functions. The claims fail to positively recite the other statutory class (machine or apparatus) by not identifying the machine or apparatus that accomplishes the method steps. The steps

might imply that a machine or apparatus is being used, but the steps do not inherently require the machine or apparatus. Therefore, the method is not a patent eligible process under 35 U.S.C. 101. See Federal Circuit Court Decision, In re Bilski, Appeal No. 2007-1130.

Claim Rejections - 35 USC § 102

3. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

(e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

4. Claims 1, 5-9 and 14-17 are rejected under 35 U.S.C. 102(e) as being anticipated by Mangipudi, et al (2004/0162901).

Regarding claims 1 and 17, Mangipudi discloses a method of processing a request to at least one server, and an article of manufacture for processing a request to at least one server, comprising a computer readable medium containing one or more programs which when executed implement the steps of:

receiving the request (**para. 45; router receiving a client request for web content**); and

scheduling submission of the request to the at least one server (**para. 47; all client requests are routed to the server selected as the most available and/or efficient server within each class according to a selected load balancing algorithm**) based on: (i) a quality-of-service (QoS) class assigned to a client from which

the request originated (**Fig. 8E, 9; para. 45, 69, 70; Class of Service (ie class) is implemented as a function of the user; user is authenticated and the respective class is assigned**); (ii) a response target associated with the QoS class (**para. 21, 25, 26, 54; routing by class to meet user's expectations, SLA metrics such as response times fall within committed levels**); and (iii) an estimated response time associated with the at least one server (**para. 55; server attributes such as response times of back-end servers are reported to router as input to policy engine**).

Regarding claim 5, Mangipudi further discloses further comprising the step of assigning the response target to the QoS class (**para. 54**).

Regarding claim 6, Mangipudi further discloses wherein the step of assigning the response target to the QoS class further comprises the step of assigning a response time target to the QoS class. (**para. 54**)

Regarding claim 7, Mangipudi further discloses wherein the step of assigning the response target to the QoS class further comprises the step of assigning a response percentile target to the QoS class. (**para. 48**)

Regarding claim 8, Mangipudi further discloses further comprising the step of estimating the response time associated with the at least one server based on one or more requests sent to the at least one server within a given time period. (**para. 55, 56**)

Regarding claim 9, Mangipudi further discloses further comprising the step of assigning a target response time to a plurality of QoS classes in which lower quality classes are assigned larger response times than higher quality classes. (**para. 38**)

Regarding claim 14, Mangipudi further discloses an apparatus for processing a request to at least one server, comprising:

a memory; **(para. 39)** and

at least one processor coupled to the memory **(para. 39)** and operative to perform the method of claim 1, and is therefore rejected under the same reason set forth in the rejection of claim 1.

Regarding claim 15, Mangipudi further discloses wherein the memory and the at least one processor form a scheduler **(Fig. 3, element 200)** that is external to the at least one server **(Fig. 3, elements 206)**.

Regarding claim 16, Mangipudi further discloses wherein the scheduler is a front-end scheduler and the at least one server is a back-end server **(Fig. 3, elements 200, 206; para. 39)**.

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

6. This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was

not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

7. The factual inquiries set forth in *Graham v. John Deere Co.*, 383 U.S. 1, 148 USPQ 459 (1966), that are applied for establishing a background for determining obviousness under 35 U.S.C. 103(a) are summarized as follows:

1. Determining the scope and contents of the prior art.
2. Ascertaining the differences between the prior art and the claims at issue.
3. Resolving the level of ordinary skill in the pertinent art.
4. Considering objective evidence present in the application indicating obviousness or nonobviousness.

8. Claims 2-4, 18-20 and 25 are rejected under U.S.C. 103(a) as being unpatentable over Mangipudi, et al, in view of Bender, et al (US 6,112,221), and further in view of Chen, et al (US 2003/0120705).

Regarding claim 2, Mangipudi discloses all of the subject matter as disclosed previously in this office action except for the following:

further comprising the step of withholding the request from submission to the at least one server when the request originated from a client assigned to a first QoS class to allow a request that originated from a client assigned to a second QoS class to meet a response target associated therewith.

However scheduling of requests based on priority queues according to response target is well known in the art. For example, Bender from the same or similar fields of endeavor discloses a server which employs a pre-emptive setting not continuously processing a request, but scheduling them according to an earliest deadline first

methodology, by alternately processing the request with the earliest deadline first, followed by that with the next earliest deadline, and so on (**col. 4, lines 52-58; col. 5, lines 27-35**). That is, in Bender, requests are scheduled based on a response target, when it is to be completed. Additionally Chen from analogous art discloses request scheduling using priority queues, in which a request in priority two queue is processed after a request in priority one is. (**fig. 1A, elements 52, 56; para. 19-21**) Thus it would have been obvious to the person of ordinary skill in the art at the time of the invention to combine the request processing method of Mangipudi with the pre-emptive scheduling method based on response target of Bender and the priority queues of Chen by scheduling requests into priority queues according to their target times. The motivation for such a combination would have been to ensure proper order of processing requests.

Regarding claims 3 and 4, the combination of Mangipudi, Bender and Chen discloses substantially all of the subject matter as disclosed previously in this office action. Mangipudi further discloses:

determining a throughput of the at least one server (para. 57), as recited in claim 3; and

monitoring a throughput of the at least one server (para. 57), as recited in claim 4.

Mangipudi does not explicitly disclose:

reducing a request withhold rate to increase throughput of the at least one server, as recited in claim 3;

varying a request withhold rate to balance the throughput and request response times, as recited in claim 4.

Bender further discloses a server which employs a pre-emptive setting of scheduling requests according to an earliest deadline first methodology (**col. 4, lines 52-58; col. 5, lines 27-35**), calculates processing time and dead line for each request (**Fig. 2, element 102**), and continues adjusting estimated processing time (**Fig. 2, element 112**).

Thus it would have been obvious to the person of ordinary skill in the art at the time of the invention to combine the request processing method of Mangipudi with the pre-emptive scheduling method based on response target of Bender and the priority queues of Chen by scheduling requests into priority queues according to their target times, monitoring the throughput of the servers, and adjusting the pre-emption rate so that the request can be processed within the target response time. The motivation for such a combination would have been to ensure proper order of processing requests.

Regarding claims 18 and 25, Mangipudi discloses assigning at least one client to a quality-of-service (QoS) class from among at least two QoS classes (**para. 70**), assigning a response target to at least one QoS class (**para. 54**); and estimating at least one response time of the at least one server based on one or more requests sent to the server within a given time period (**para. 55, 56**)

Mangipudi discloses all of the subject matter except:

withholding requests associated with a first one of the at least two QoS classes to allow requests associated with a second one of the at least two QoS classes to meet its response target based on the at least one estimated response time.

However scheduling of requests based on priority queues according to response target is well known in the art. For example, Bender from the same or similar fields of endeavor discloses a server which employs a pre-emptive setting not continuously processing a request, but scheduling them according to an earliest deadline first methodology, by alternately processing the request with the earliest deadline first, followed by that with the next earliest deadline, and so on (**col. 4, lines 52-58; col. 5, lines 27-35**). That is, in Bender, requests are scheduled based on a response target, when it is to be completed. Additionally Chen from analogous art discloses request scheduling using priority queues, in which a request in priority two queue is processed after a request in priority one is. (**fig. 1A, elements 52, 56; para. 19-21**) Thus it would have been obvious to the person of ordinary skill in the art at the time of the invention to combine the request processing method of Mangipudi with the pre-emptive scheduling method based on response target of Bender and the priority queues of Chen by scheduling requests into priority queues according to their target times. The motivation for such a combination would have been to ensure proper order of processing requests.

Claims 19 and 20 are substantial duplicates of claims 3 and 4, respectively, and are therefore rejected under the same reason set forth in the rejection of claims 3 and 4, respectively.

9. Claim 8 is rejected, in an alternative, under U.S.C. 103(a) as being unpatentable over Mangipudi, et al in view of Veres, et al (US 6,807,156).

Regarding claim 8, Mangipudi discloses all of the subject matter as previously recited in this office action. Furthermore, Veres from the same or similar fields of endeavor discloses further comprising the step of estimating the response time (**col. 13, lines 46-47**) associated with the at least one server or applications based on one or more requests sent to the at least one server or applications within a given time period (**time window of measurement as shown in Fig. 2; col. 13, lines 36-47**). Thus it would have been obvious to the person of ordinary skill in the art at the time of the invention to combine the request processing method of Mangipudi with the response time estimating method of Veres by periodically sending requests to the applications and servers to estimate the response time. The motivation for such a combination would have been to ensure service level agreement based on response time is met.

10. Claims 10-12 are rejected under U.S.C. 103(a) as being unpatentable over Mangipudi, et al in view of Menditto, et al (US 6,981,029).

Regarding claim 10, Mangipudi discloses all of the subject matter as disclosed previously in this office action except for the following:

determining dispatch times for requests from a difference between at least one predicted response time of the at least one server and the target response time corresponding to the QoS class of the request; and
sending requests to the at least one server based on dispatch times.

However Mangipudi discloses using response time as a metric to select a server to meet SLA commitments (**para. 61, 62**). Menditto from the same or similar fields of endeavor discloses a content gateway making routing decisions based on the request, selecting a server satisfying the request and depending on various factors such as server load, and forwarding the request to the selected server (**col. 3, lines 11-61**). Thus it would have been obvious to the person of ordinary skill in the art at the time of the invention to combine the request processing method of Mangipudi with the QoS enforcement approach by Menditto by selecting a server that can timely process the request. The motivation for such a combination would have been to ensure service level agreement based on response time is met.

Regarding claim 11, Mangipudi further discloses:

wherein a plurality of applications are running on the at least one server and requests are routed to applications (**Fig. 3, application servers 216; para. 45**), further comprising the steps of:

estimating response times of applications based on one or more requests sent to the applications within a time period. (**para. 55, 56**)

Mangipudi does not explicitly disclose:

sending a request to an application whose estimated response time is not greater than a target response time corresponding to the QoS class of the request.

However Mangipudi discloses using response time as a metric to select a server to meet SLA commitments (**para. 61, 62**). Menditto from the same or similar fields of endeavor discloses selecting an optimal server based on a set of rules, defining as

producing the quickest response time to the request. (**col. 6, lines 16-40**) Thus it would have been obvious to the person of ordinary skill in the art at the time of the invention to combine the request processing method of Mangipudi with the QoS enforcement approach by Menditto by selecting a server that can timely process the request. The motivation for such a combination would have been to ensure service level agreement based on response time is met.

Regarding claim 12, Mangipudi discloses all of the subject matter as recited above except:

further comprising the step of varying a number of requests sent to applications so that estimated response times of applications are not greater than target response times of QoS classes corresponding to requests sent to the applications.

However Mangipudi discloses sending requests to applications to a server based on response time (**para. 61, 62**) Menditto from the same or similar fields of endeavor discloses a content gateway updating policies regarding processing of requests based on service level agreements (**col. 7, lines 1-52**). Thus it would have been obvious to the person of ordinary skill in the art at the time of the invention to combine the request processing method of Mangipudi with the QoS enforcement approach by Menditto by selecting a server that can timely process the request. The motivation for such a combination would have been to ensure service level agreement based on response time is met.

11. Claim 13 is rejected under 35 U.S.C. 103(a) as being unpatentable over Mangipudi, et al in view of Menditto, et al, and further in view of Lu, et al (US 6,772,211).

Regarding claim 13, the combination of Mangipudi and Menditto discloses all of the subject matter as disclosed previously in this office action except *wherein the at least one server comprises a plurality of servers and each application runs on a different one of the plurality of servers*.

Lu from the same or similar fields of endeavor discloses methods to switch client packets to one server among a group of servers (**col. 4, lines 50-53**) and applications have their own dedicated servers (**col. 5, lines 24-26**).

Thus it would have been obvious to the person of ordinary skill in the art at the time of the invention to combine the packet processing method of Mangipudi and Menditto with the servers and applications of Lu by implementing the method and each application on a separate server. The motivation for such a combination would have been to modularize the features for scalability and performance.

12. Claims 21-23 are rejected under U.S.C. 103(a) as being unpatentable over Mangipudi, et al in view of Bender, et al, Chen, et al and Menditto et al.

Claims 21-23 are substantial duplicates of claims 10-12, respectively, and are therefore rejected under the same reason set forth in the rejection of claims 10-12, respectively.

13. Claim 24 is rejected under U.S.C. 103(a) as being unpatentable over Mangipudi, et al in view of Bender, et al, Chen, et al and Menditto et al, and in further view of Lu, et al.

Claim 24 is a substantial duplicate of claims 13 and is therefore rejected under the same reason set forth in the rejection of claim 13.

Conclusion

14. Any inquiry concerning this communication or earlier communications from the examiner should be directed to LUAT PHUNG whose telephone number is (571) 270-3126. The examiner can normally be reached on M-Th 7:30 AM - 5:00 PM, F 7:30 AM - 4:00 PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Ricky Q. Ngo can be reached on 571-272-3139. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/L. P./

Examiner, Art Unit 2416

/Ricky Ngo/

Supervisory Patent Examiner, Art Unit 2416